

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:52:44 ; Search time 170.72 Seconds
(without alignments)
15.909 Million cell updates/sec

Title: US-09-331-631a-8_COPY_80_119
Perfect score: 225
Sequence: 1 PEDPQRREYECQECRQOEERQOPQCQQRCLKRFEGEQQQ 40

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	221	98.2	588	1 FWCNAB	alpha-globulin B p
2	213	94.7	509	2 S08059	alpha-globulin typ
3	142	63.1	605	2 S06398	alpha-globulin typ
4	119	52.9	566	2 S22477	vicilin precursor
5	86	38.2	1905	2 T18267	multidrug resistanc
6	83	36.9	429	2 S29565	apolipoprotein A-I
7	83	36.9	1403	2 S24548	homeotic protein p
8	82	36.4	339	1 TWHU2D	transcription init
9	82	36.4	401	2 A47141	apolipoprotein A-I
10	80	35.6	1457	2 T14577	protein kinase Yak
11	78	34.7	613	2 S27770	hypothetical prote
12	78	34.7	648	1 Q01150	protein kinase (EC
13	77	34.2	425	2 T18592	hypothetical prote
14	77	34.2	758	2 S54522	hypothetical prote
15	77	34.2	1761	2 T13675	hypothetical prote
16	76.5	34.0	388	2 T31887	hypothetical prote
17	76.5	34.0	388	2 T31888	hypothetical prote
18	76.5	34.0	467	1 A49377	involucrin - mouse
19	76.5	34.0	738	2 S37876	glutamine-rich pro
20	76	33.8	572	2 T29880	hypothetical prote
21	76	33.8	600	2 T18593	hypothetical prote
22	76	33.8	905	1 RGRY55	regulatory protein
23	76	33.6	1038	2 T02634	rep protein homolo
24	75.5	33.6	438	2 T31869	hypothetical prote
25	75.5	33.6	445	2 T31898	hypothetical prote
26	75	33.3	1023	2 T13068	hypothetical prote
27	75	33.3	1027	2 T13071	CLOCK protein - fr
28	75	33.3	1154	2 S69206	regulator protein
29	75	33.3	1969	2 T08875	histidine kinase h

30	74.5	33.1	1390	2 T14004	trfA protein - sli
31	73.5	32.7	411	2 T29475	hypothetical prote
32	73	32.4	139	2 A26892	Mopa box protein -
33	73	32.4	646	2 D82493	conserved hypochet
34	73	32.4	1505	2 JC4851	hypoxia-inducible
35	72.5	32.2	292	2 B22364	alpha/beta-gliadin
36	72.5	32.2	307	2 S10015	alpha/beta-gliadin
37	72.5	32.2	319	2 D22364	alpha/beta-gliadin
38	72.5	32.2	326	2 T00377	alpha/beta-gliadin
39	72.5	32.2	1069	2 T00377	KIAA0642 protein -
40	72	32.0	538	2 S29521	casein kinase I ho
41	72	32.0	4957	2 T03455	ALR protein - huma
42	72	32.0	5262	2 T03454	ALR protein - huma
43	71	31.6	286	1 EEMTA	alpha/beta-gliadin
44	71	31.6	286	2 S07923	alpha/beta-gliadin
45	71	31.6	1365	2 S14871	suppressor two of

ALIGNMENTS

RESULT 1

FWCNAB

N:Alternate names: seed storage protein; vicilin precursor

C:Species: Gossypium hirsutum (upland cotton)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999

C:Accession: A30838; S06911

R:Chlan, C.A.; Pyle, J.B.; Legocki, A.B.; Dure III, L.

Plant Mol. Biol. 9, 475-489, 1986

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination XVIII

A:Reference number: A30838

A:Molecule type: mRNA

A:Residues: 1-588 <CHL>

A:Cross-references: GB:M16891; NID:G167374; PIDN:AAA33071.1; PID:G167375

A:Experimental source: var. Coker 201

R:Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.

Plant Mol. Biol. 9, 533-546, 1987

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.

A:Reference number: S06398

A:Accession: S06911

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-81 <CH2>

C:Comment: This is a seed storage protein.

C:Superfamily: glycinin

C:Keywords: glycoprotein; seed; storage protein

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-588/Product: alpha-globulin storage protein #status predicted <MAN>

F:417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 98.2% Score 221, DB 1, Length 588;

Best Local Similarity 97.5% Pred. No. 5e-16;

Matches 39; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PEDPQRREYECQECRQOEERQOPQCQQRCLKRFEGEQQQ 40
DB 80 PEDPQRREYECQECRQOEERQOPQCQQRCLKRFEGEQQQ 119

RESULT 2

S08059

N:Alternate names: seed storage protein

C:Species: Gossypium hirsutum (upland cotton)

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993

C:Accession: S08059

R:Chlan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.

Plant Mol. Biol. 9, 533-546, 1987

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.

A:Reference number: S06398

C:Species: *Drosophila melanogaster* (wrosophila melanogaster)
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 15-Oct-1999
C:Accession: S24548; J01397; A41089
R:Doe, C.O.
submitted to the EMBL Data Library, February 1992
A:Reference number: S24548

A:Accession: S24548
A:Molecule type: DNA
A:Residues: 1-1403 <DOE>
A:Cross-references: EMBL:Z11743; NID:g8383; PID:g8384
R:Matsuzaki, F.; Koizumi, K.; Hama, C.; Yoshioke, T.; Nabeshima, Y.
Biochem. Biophys. Res. Commun. 182, 1326-1332, 1992
A:Title: Cloning of the *Drosophila* prospero gene and its expression in ganglion mother cells
A:Reference number: J01397; MUID:92171948
A:Accession: J01397
A:Molecule type: mRNA
A:Residues: 1-417, 'Q', 419-1403 <MAT>
A:Cross-references: GB:D10609; DDBJ:D90560; NID:g217345; PID:d1001939; PID:g217346
R:Vaesslin, H.; Grell, E.; Wolff, E.; Bieri, E.; Jan, L.Y.; Jan, Y.N.
Cell 67, 941-953, 1991
A:Title: prospero is expressed in neuronal precursors and encodes a nuclear protein that is required for the development of the CNS
A:Reference number: M1089; MUID:92069760
A:Accession: M1089
A:Molecule type: mRNA
A:Residues: 1-75, 'GQDAERAVMPDDEAGPCRNEMPA', 99-119, 'NLAIQFVQVAAAAAITTALLPIPG', 145-676, 'C
A:Cross-references: GB:M6189; NID:g158183; PID:g158184
C:Genetics:
A:Gene: FlyBase:pros
A:Cross-references: FlyBase:FBgn0004595
C:Keywords: DNA binding; nucleus; transcription regulation
F:717-734/Region: glutamine-rich
F:991-998/Region: nuclear location signal
F:1029-1049/Region: glutamine-rich

```

Query Match Similarity 36.9% Score 83 DB 2 Length 1403;
Best Local Similarity 54.8% Pred No 0.23;
Matches 17; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 10 ECQCECRQDEERKQPPCCQRCIKRFEQEQQQ 40
  |||:::|||||
716 EQQQQQQQQQQQQQQQQQQQQEQQRFEQEQEQ 746

```

RESULT 8
TMW02D
transcription initiation factor IID - human
N:Alternate names: TATA-binding protein
C:Species: Homo sapiens (man)
C:Date: 20-Jul-1990 #sequence_revision 19-May-1995 #text_change 18-Feb-2000
A:Accession: A34830; A34831; S10944; I60128
R:Peterson, M.G.; Tanese, N.; Pugh, B.F.; Tjian, R.
Science 248, 1625-1630, 1990
A:Title: Functional domains and upstream activation properties of cloned human TATA bind
A:Reference number: A34830; MUID:90302006
A:Accession: A34830
A:Molecule type: mRNA
A:Residues: 1-339 <PEP>
A:Cross-references: GB:M5654; NID:g339491; PIDN:AAA36731.1; PID:g339492
R:Kao, C.C.; Lieberman, P.M.; Schmidt, M.C.; Zhou, Q.; Pei, R.; Berk, A.J.
Science 248, 1645-1649, 1990
A:Title: Cloning of a transcriptionally active human TATA binding factor.
A:Reference number: A34831; MUID:90302010
A:Accession: A34831
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-17, 'N', '19-186, 'R', 188-339 <KAO>
R:Horikoshi, A.; Sinn, E.; Yamamoto, T.; Wang, J.; Roy, A.; Horikoshi, M.; Koeder, R.G.
Nature 346, 387-390, 1990
A:Title: Highly conserved core domain and unique N terminus with presumptive regulatory
A:Reference number: S10944; MUID:90326195
A:Accession: S10944
A:Molecule type: mRNA
A:Residues: 1-91, 96-339 <HOF>
A:Cross-references: EMBL:X54993; NID:g37065; PIDN:CAA8736.1; PID:g37066
R:Kao, C.; Lieberman, P.; Schmidt, M.; Zhou, Q.; Pei, R.; Berk, A.J.
Science 248, 1626-1690
A:Title: Cloning of the human TATA binding factor: Expression of a transcriptionally act
A:Reference number: I60128

A:Accession:160128
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-186, 'R', 188-299, 'MIKPR', 300-339 <RES>
A:Cross-references: GB:M34960; NID:g339493; PID:g339494
C:Genetics:
A:Gene: GDB:TBPI; GTF2D1
A:Cross-references: GDB:138768; OMIM:600075
A:Map position: 6927-6927
C:Superfamily: human transcription initiation factor IID
C:Keywords: alternative splicing; DNA binding; nucleus; transcription initiation
C:5-95/Region: glutamine-rich

	36.4%;	Score	82;	DB	1;	Length	339;
Query Match	41.0%;	Pred.	No.	0.09;			
Best Local Similarity							
Matches	16;	Conservative	14;	Mismatches	9;	Indels	0;
Gaps							0;
OY	2	EDPQRRFECCOECNOEERDOPQCQRCLKRFEBDEBOOO	40				
		: : : : : : : : : : : : : : : : : : : :					
db	53	EEOGROROOOOOOOOOOOOOOOOOOOOOOOOOOOOOOOO	91				

```

RESULT      9
A47141
apolipoprotein A-IV I isoform - baboon (fragment)
C:Species: Papio sp. (baboon)
C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 12-Apr-1995
C:Accession: A47141
R:Hiixon, J.E.; Kummerer, G.M.; Mott, G.E.; Britten, M.L.; Birnbaum, S.; Powers, P.K.
J. Biol. Chem. 268, 15667-15673, 1993
A:Title: Baboon apolipoprotein A-IV. Identification of Lys76-->Glu that distinguishes
A:Reference number: A47141; MUID:93340170
A:Accession: A47141
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-401 <HTX>
A:Experimental source: intestine
A:Note: Sequence extracted from NCBI backbone (NCBIN:136009, NCBIIP:136010)
;Superfamily: apolipoprotein A-I

```

[illegible]

```

RESULT 10
T14577
protein kinase Yaka (EC 2.7.1.-) - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14577
R:Kuspa, A.; Lu, S.; Souza, G.M.
submitted to the EMBL Data Library, January 1998
A:Description: Yaka, a protein kinase required for the growth to development transiti
A:Reference number: Z18146
A:Accession: T14577
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1457 <KUS>
A:Cross-References: EMBL:AF045453; NID:g2854116; PID:g2854117; PIDN:AAC02554.1
C:Genetics:
A:Gene: yaka
C:Keywords: ATP; phosphoprotein; phosphotransferase; serine/threonine-specific protei

```

A/status: preliminary; translated from GB/EMBL/DBJ
A/molecule type: DNA

A:Residues: 1-1761 <BOL>
A:Cross-references: EMBL:AL031130; NID:e1316407; PTD:e1316410; PIDN:CAA20016.1
C:Genetics:
A:Introns: 143/3; 237/3; 280/3
A:Note: EG:EG0002.3

Query Match 34.28; Score 77; DB 2; Length 1761;
Best Local Similarity 42.18; Pred. No. 1.2;
Matches 16; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 3 DPQRRECCQCECHQCEERQDPCCQRCCKRFEQEQQ 40
111 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 : 11 :
Db 1480 DMORYVCRMD000000000000000000000000 1517

Search completed: March 1, 2001, 15:52:45
Job time: 569 sec

